Computational Center for Biomolecular Complexes
C^2BC

LEADERSHIP TEAM

Dr. Wah Chiu
Baylor College of Medicine
wah@bcm.tmc.edu

Dr. Helen Berman
berman@rcsb.rutgers.edu
The State University of New Jersey at Rutgers

Dr. Chandrajit Bajaj
bajaj@cs.utexas.edu
University of Texas at Austin

Dr. Arthur Olson
olson@scripps.edu
The Scripps Research Institute

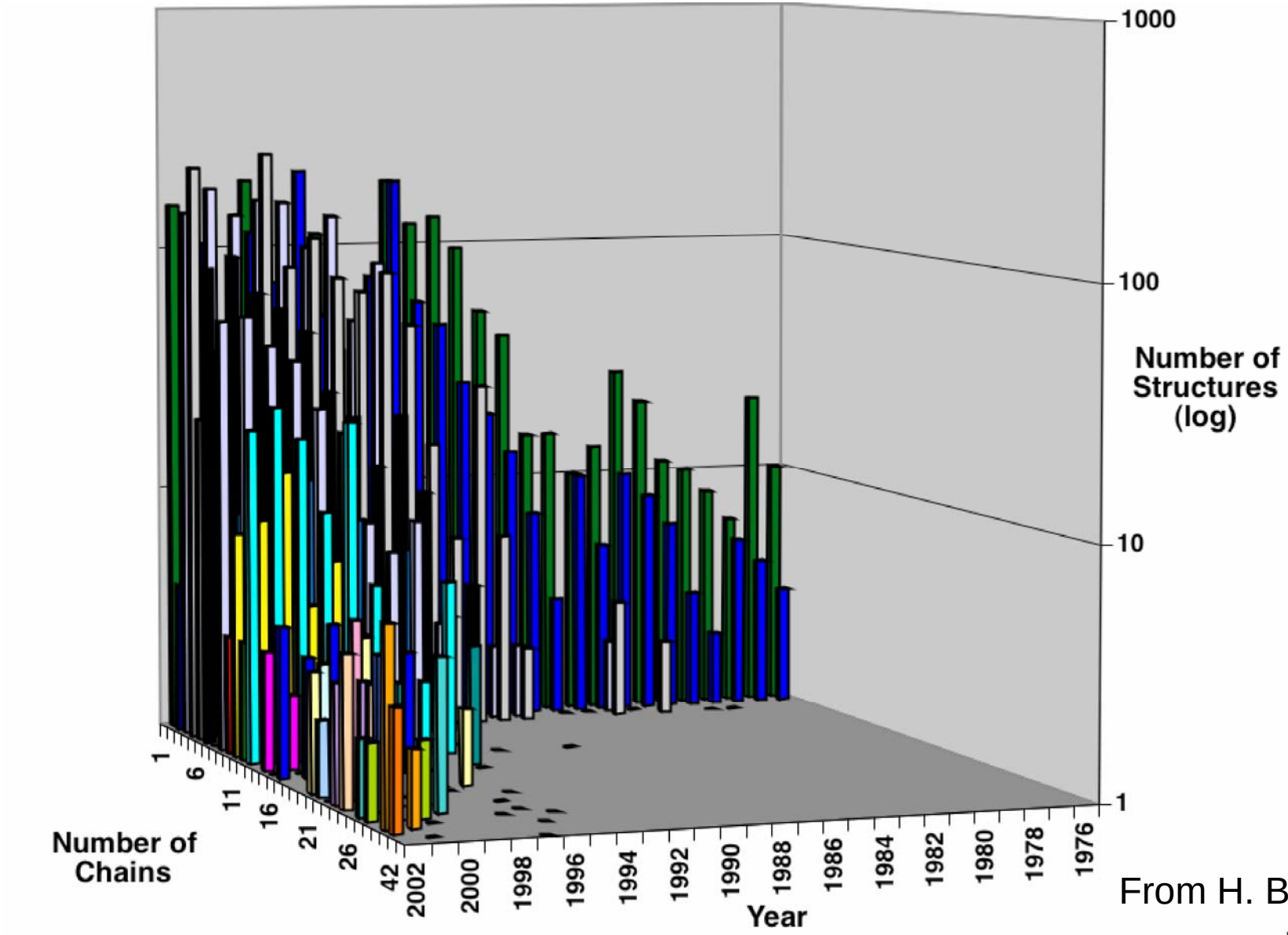
http://ncmi.bcm.tmc.edu/ncmi/ccbc

National Center for Research Resources

About C^2BC

- Focused on the development of computational tools for studying functional mechanisms of biological complexes both in vitro and inside a cell
- A virtual center of interdisciplinary research across institutions

Trend in PDB Structures

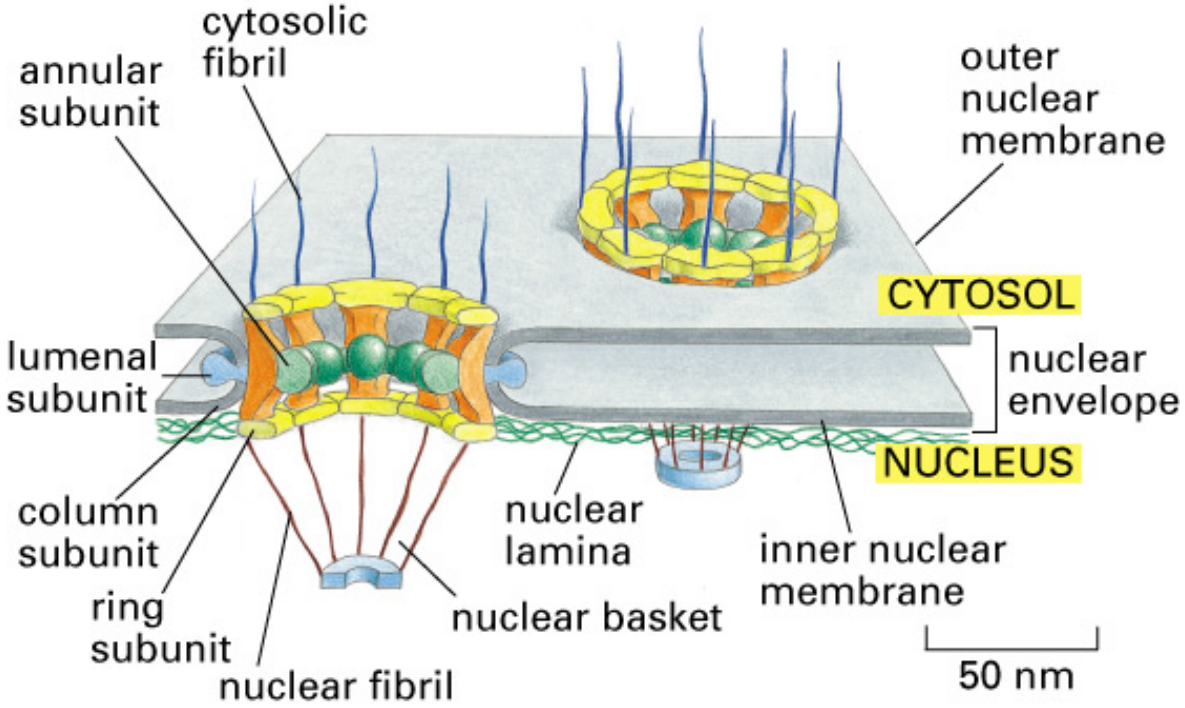


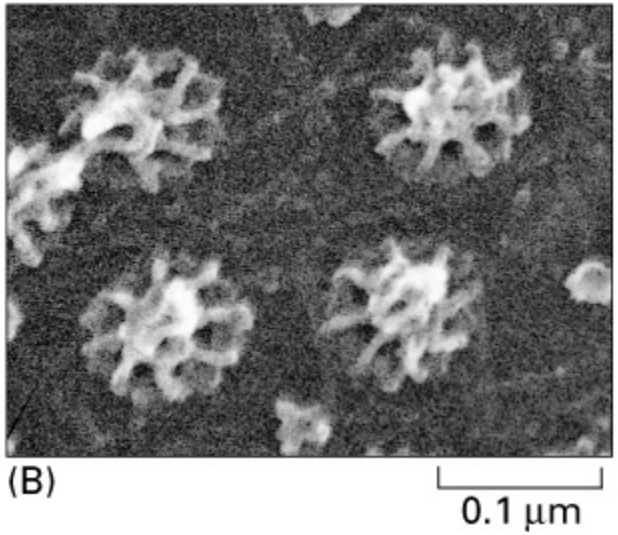
From H. Berman

Why Study Large Complexes?

- Proteins typically function in association with other proteins.
- Protein complexes are important for virtually every biological process and most diseases.
- Genome sequences identify tens of thousands of genes; linking these to 200-300 core biological processes will make their study manageable.
- Recently developed and/or improved technologies and methodologies make studies of large complexes more feasible and informative.

Example: Nuclear Pore (152 MDa)

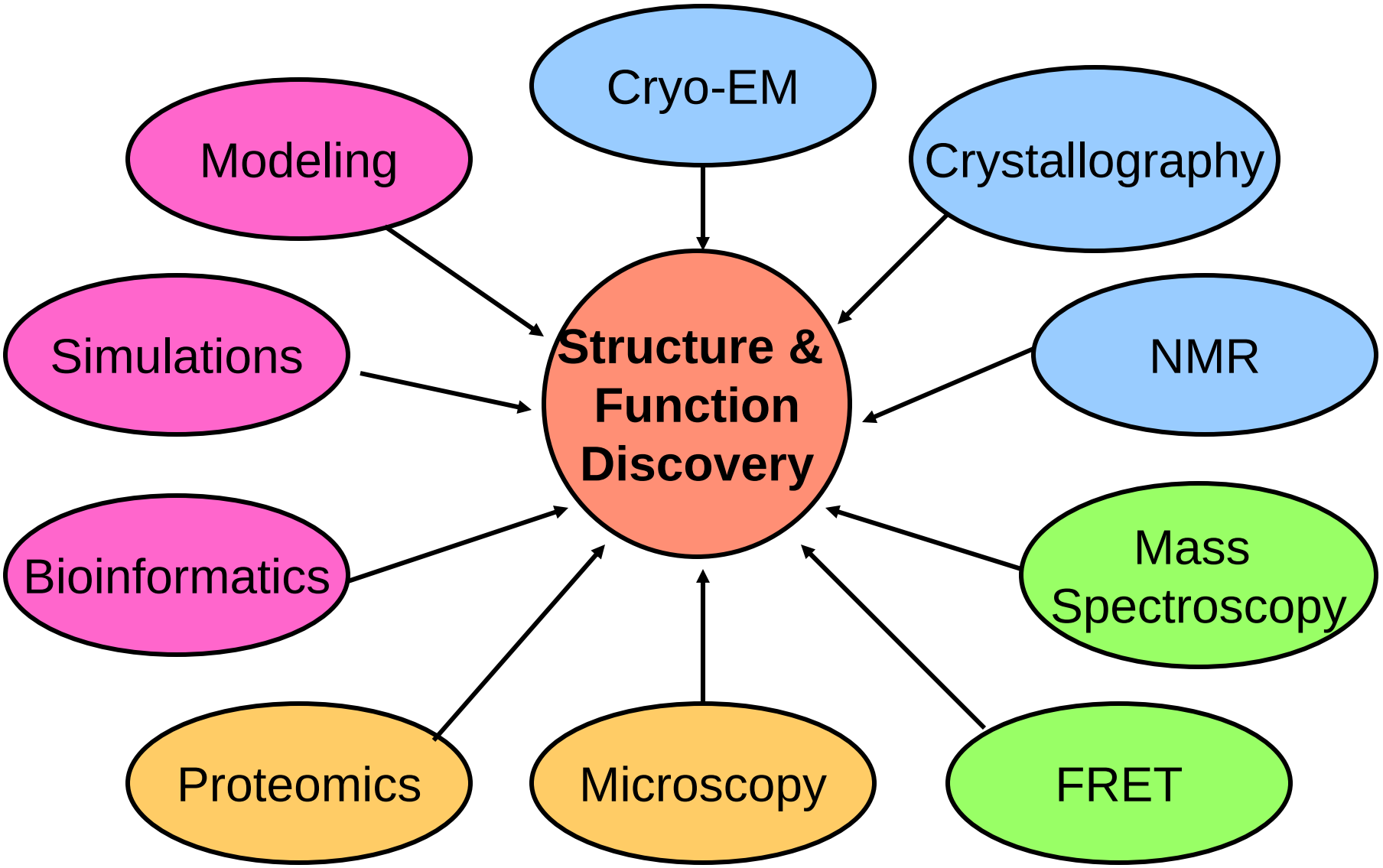




Over 50 individual protein components

From Molecular Biology of the Cell

Approaches for Studying Complexes

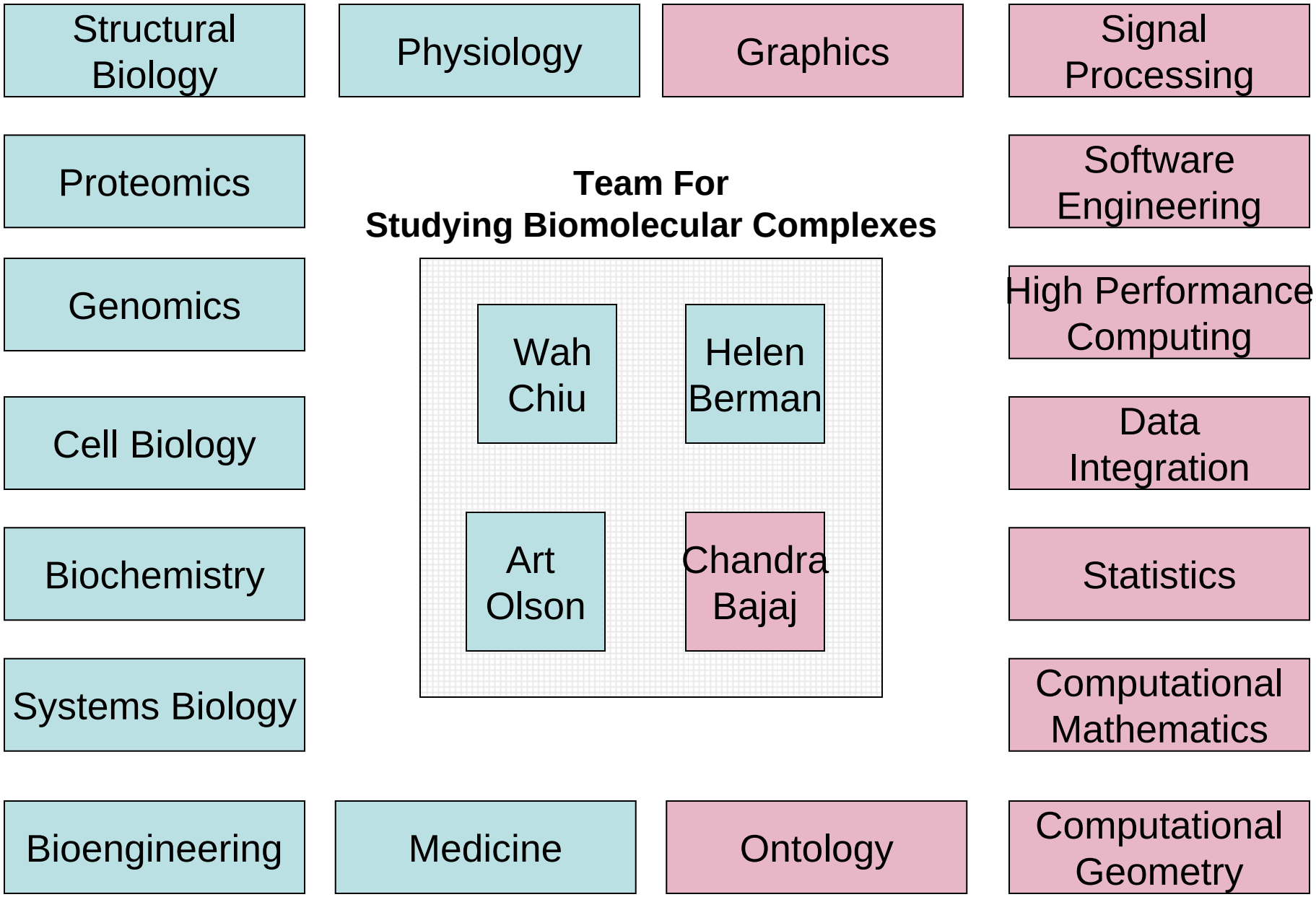


Experimental Challenges

- Identification of complexes (transient vs. persistent)
- Purification of complexes
- Sample quantity/concentration
- Multiple functional states
- Validation of complexes in the living cell
- Spatial and temporal location
- Improvements of existing methodologies
- High throughput

Computational Challenges

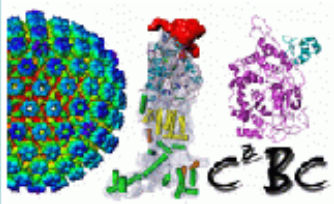
- Large data set management
- Structures of multiple conformational states
- Interplay between structure refinement and model building
- Distributed, heterogeneous data types
- Visualization
- Annotation
- Data mining
- Structure validation in cell
- Archival of data accessible to non-experts
- Specialized algorithm and user-friendly software


Structural Biology
Physiology
Graphics
Signal Processing
Proteomics
Team For
Studying Biomolecular Complexes
Software Engineering
Genomics
Wah Chiu
Helen Berman
High Performance Computing
Cell Biology
Data Integration
Biochemistry
Art Olson
Chandra Bajaj
Statistics
Systems Biology
Computational Mathematics
Bioengineering
Medicine
Ontology
Computational Geometry

Approach to Organize the Center

- Select appropriate biological model driving projects
- Define the computational challenges, gaps and opportunities through workshops with participants from biomedicine and computational sciences
- Personal contacts with established investigators of different specialists
- Regular meetings of leadership team via internet and visits

Workshops



Computational Center for Biomolecular Complexes
C²BC

Home

Investigators

Mission

Organization

Communication

Intranet

Workshops

Workshops at C²BC

- The Feb 2006 Computational Center for Biological Complex Planning Workshop
- The Sept 2005 Workshop on Visualization of Large Biomolecular Complexes
- The May 2005 Workshop on Structural and Computational Proteomics of Biological Complexes
- The Oct 2004 Cryo-Electron Microscopy Structure Deposition Workshop
- cryo-EM Dictionary development site

Hosted by NCMI
Phone: 713-798-6989
Fax: 713-798-1625

1 Baylor Plaza,Houston,TX,77030

Privacy Notices
EMail Webmaster
last modified, Mar. 09, 2005

Grant number: P20 RR020647
NIH Program Administrator: Dr. Greg Farber
Supported by NCRR
NIH announcement

Challenges to Interdisciplinary Research

- Disincentives for interdisciplinary cooperation
  - *No credit for freely distributing things because it is hard to track*
  - *No incentive for implementation and pedestrian work*
- Interdisciplinary communication
  - *Finding technical information from another domain*
- Finding the important problems in each discipline
- Finding the appropriate publication to receive credit
- Interdisciplinary training and education
- Lack mechanisms to support the effort to establish and maintain repositories
  - *software*
  - *data*

Basic Ground Rules

- Initially identify the brightest investigators with relevant expertise
- Participants have common intellectual interests and mutual respect (i.e. equal partners)
- Create new partnerships
- Share data and ideas; encourage “your data is ours and our data is yours”
- Develop a fair credit recognition practice
- Sufficient funding and food to glue the investigators together

Potential Participants

Management Team: Wah Chiu, Chandrajit Bajaj, Helen Berman, Art Olson

Imaging
W Chiu (BCM)
P Matsudaira (MIT)

Computing
P Penczek (UTH)
C Yang (LBNL)
D Scott (Rice U)
I Dhillon (UT)
S Ludtke (BCM)
W Jiang (Purdue U)

Visualization
A Olson (Scripps)
D Goodsell (Scripps)
C Bajaj (UT Austin)
T Funkhouser (Princeton U)
J Warren (Rice U)

Data & Software Integration
M Baker (BCM)
K Henrick (EBI)
M Sanner (Scripps)

Modeling
D Baker (U Washington)
L Kavarki (Rice U)
H Wolfson (TAU)
R Nussinov (NCI)
T Ju (Washington U)

Knowledge Discovery
H Berman (Rutgers U)
C Lawson (Rutgers U)

Biological Processes: Viral infections, EC Coupling, Nuclear Transport, etc…
J King (MIT), P Chisholm (MIT), S Hamilton (BCM), I Serysheva (BCM), M Rout (Rockefeller U)

Potential Participants

Management Team: Wah Chiu, Chandrajit Bajaj, Helen Berman, Art Olson

Imaging
W Chiu (BCM)
P Matsudaira (MIT)

Computing
P Penczek (UTH)
C Yang (LBNL)
D Scott (Rice U)
I Dhillon (UT)
S Ludtke (BCM)
W Jiang (Purdue U)

Visualization
A Olson (Scripps)
D Goodsell (Scripps)
C Bajaj (UT Austin)
T Funkhouser (Princeton U)
J Warren (Rice U)

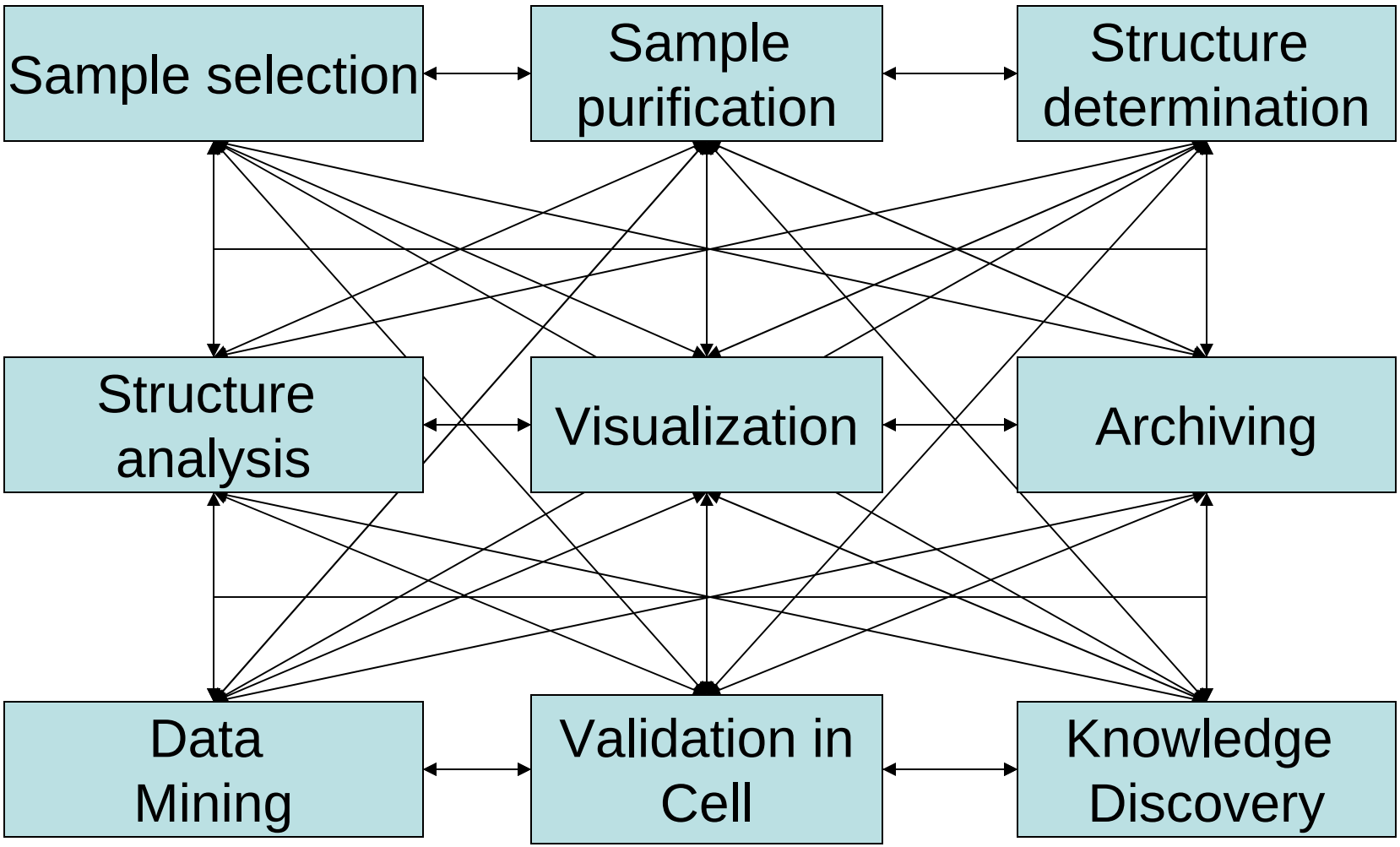
Data & Software Integration
M Baker (BCM)
K Henrick (EBI)
M Sanner (Scripps)

Modeling
D Baker (U Washington)
L Kavarki (Rice U)
H Wolfson (TAU)
R Nussinov (NCI)
T Ju (Washington U)

Knowledge Discovery
H Berman (Rutgers U)
C Lawson (Rutgers U)

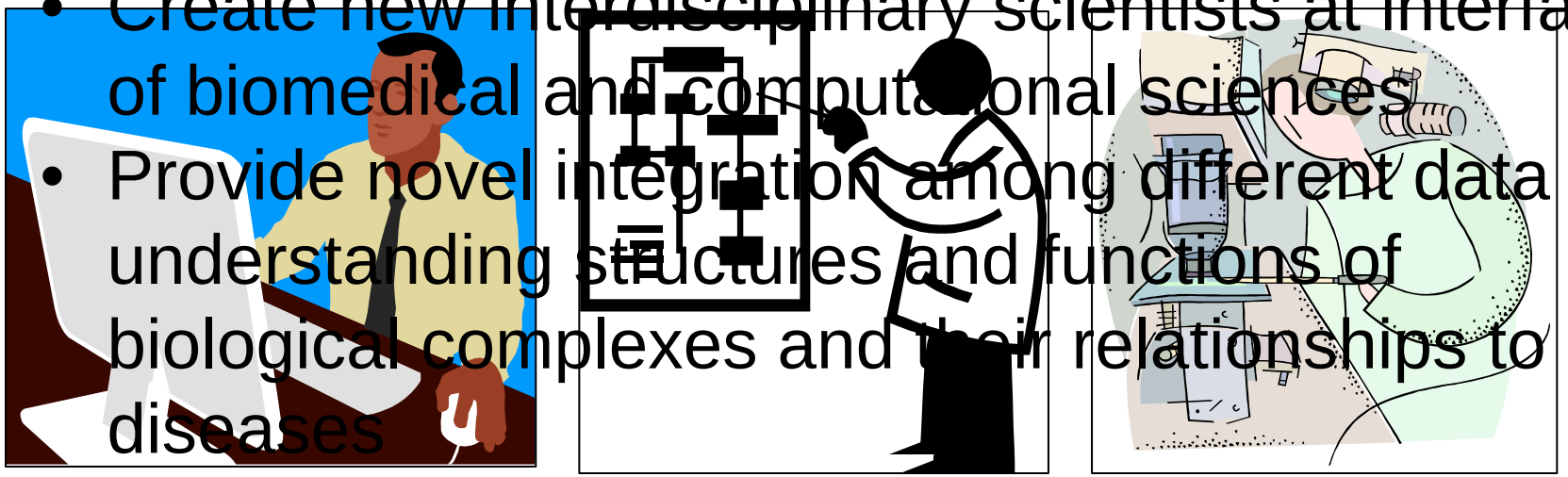
Biological Processes: Viral infections, EC Coupling, Nuclear Transport, etc…
J King (MIT), P Chisholm (MIT), S Hamilton (BCM), I Serysheva (BCM), M Rout (Rockefeller U)

Interactive and Interdisciplinary Pipeline for Studying Complexes

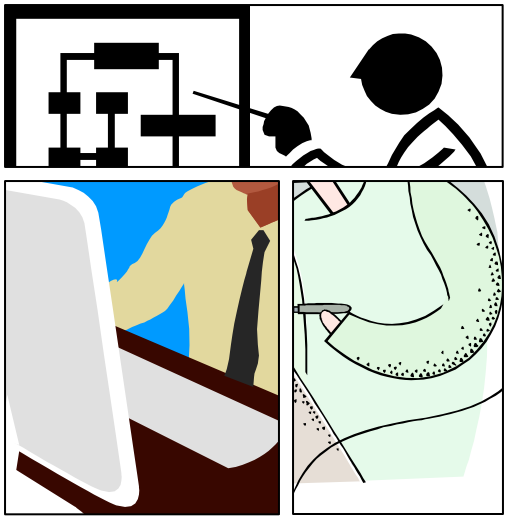


Center Outcomes

- Create new interdisciplinary scientists at interface of biomedical and computational sciences
- Provide novel integration among different data for understanding structures and functions of biological complexes and their relationships to diseases
- Generate computational toolbox freely accessible to biologists to make new discoveries



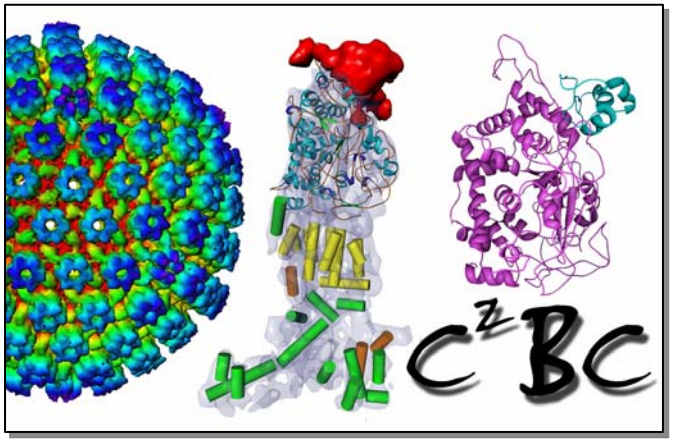
Computational Science Engineering Biomedical Science

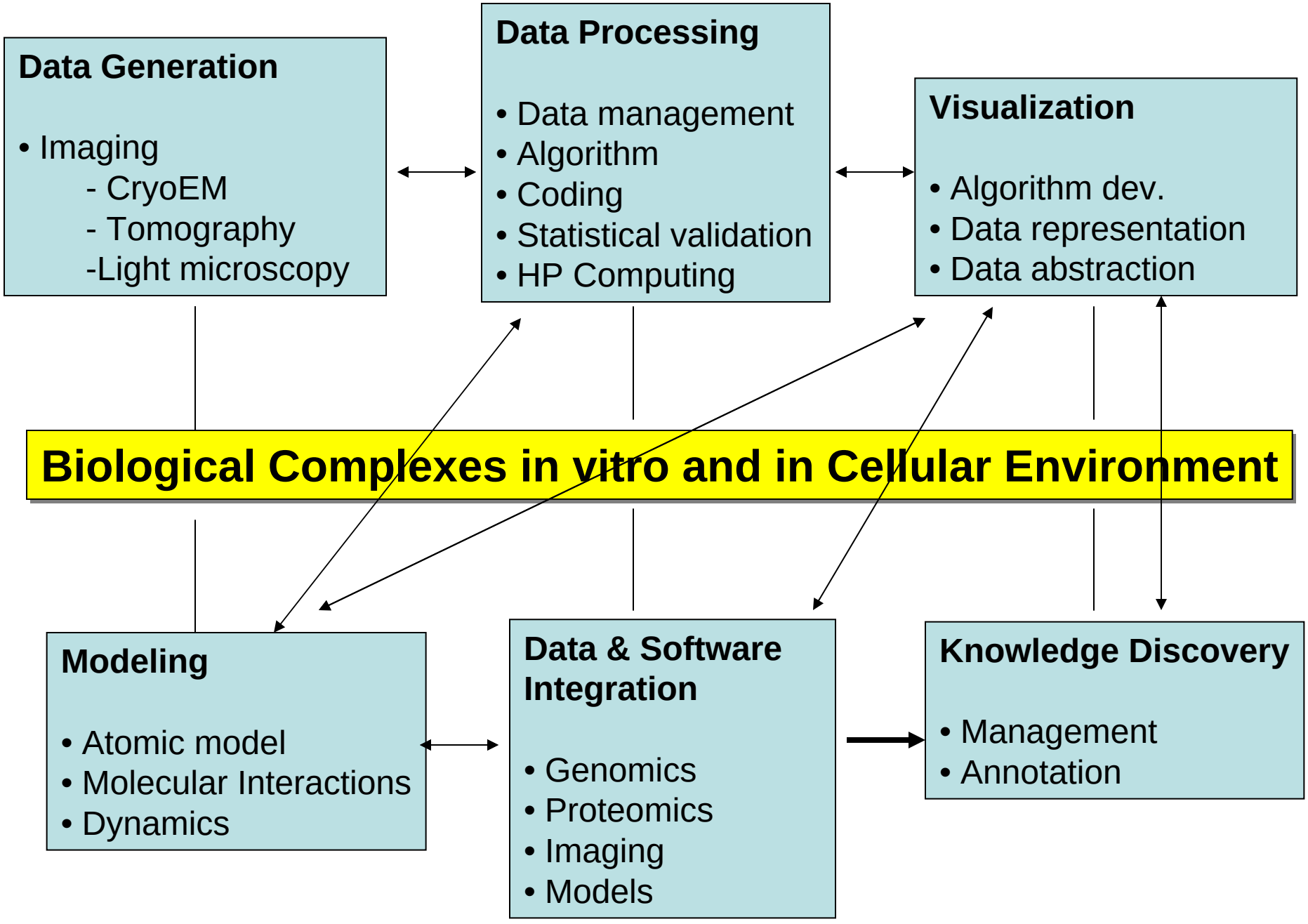


Integrated Quantitative Biomedicine

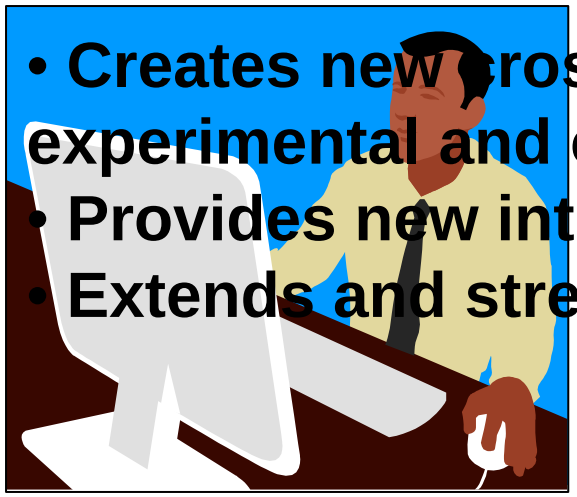
Computational Center for Biological Complexes

http://ncmi.bcm.tmc.edu/ncmi/ccbc



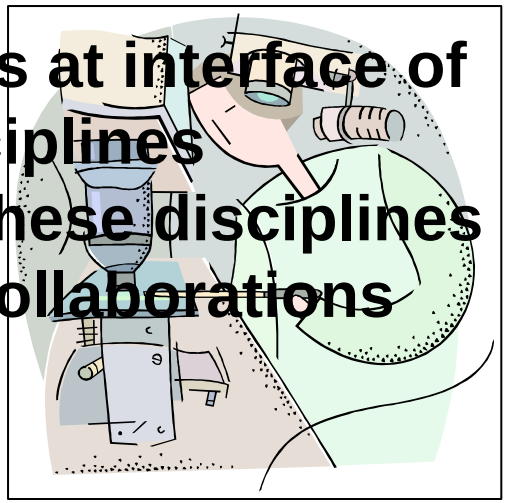

Data Generation
• Imaging
- CryoEM
- Tomography
-Light microscopy
Data Processing
• Data management
• Algorithm
• Coding
• Statistical validation
• HP Computing
Visualization
• Algorithm dev.
• Data representation
• Data abstraction
Biological Complexes in vitro and in Cellular Environment
Modeling
• Atomic model
• Molecular Interactions
• Dynamics
Data & Software Integration
• Genomics
• Proteomics
• Imaging
• Models
Knowledge Discovery
• Management
• Annotation

- **Creates new cross-trained scientists at interface of experimental and computational disciplines**
- **Provides new integration between these disciplines**
- **Extends and strengthens existing collaborations**

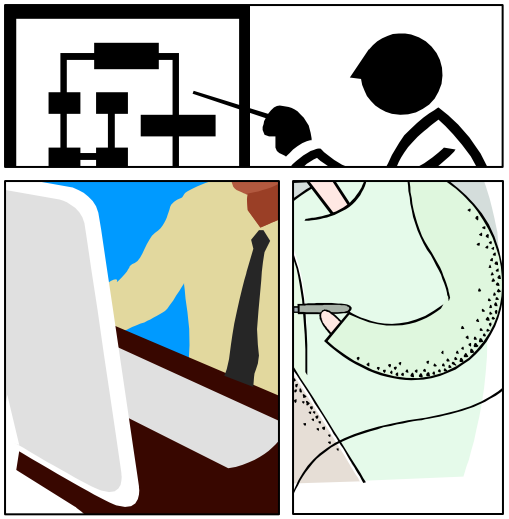


Computational Science

Engineering



Biological Science

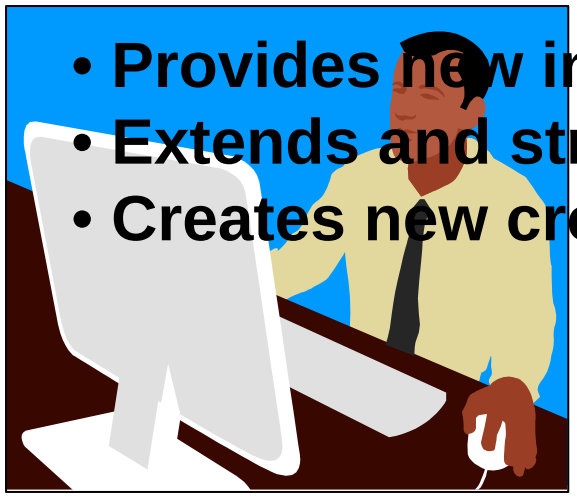


Integrated Quantitative Biomedical Science

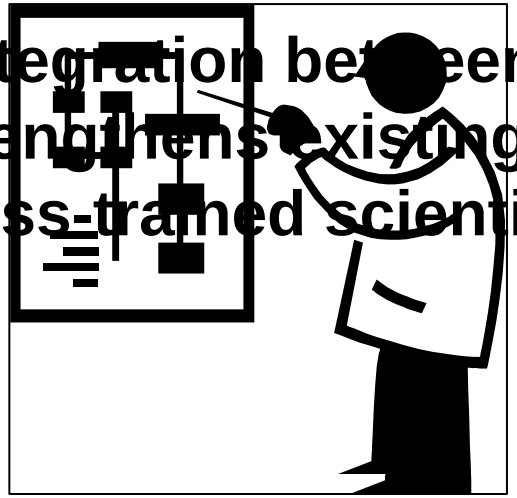
C^2BC Theme and Policy

- Computational methodology innovations
- Establishing standards
- Computational methodology validation
- Cellular validation
- Adopting open source policy
- Community participation
- Enabling tools for biological end-users

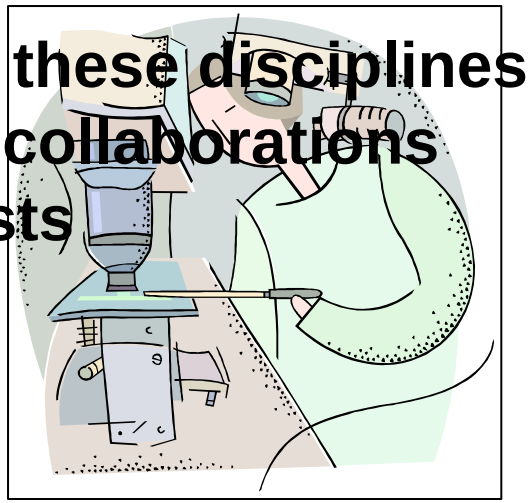
- **Provides new integration between these disciplines**
- **Extends and strengthens existing collaborations**
- **Creates new cross-trained scientists**



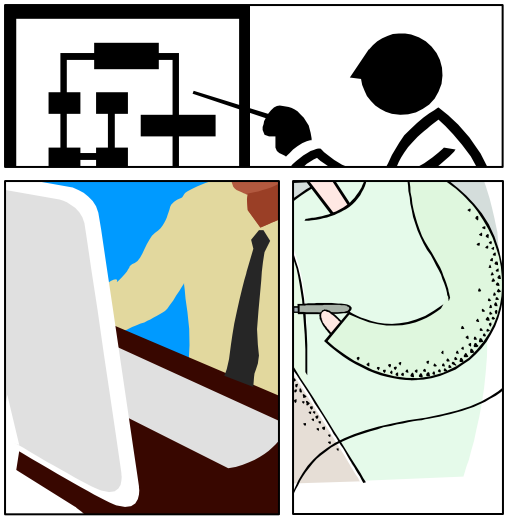
Computational Science



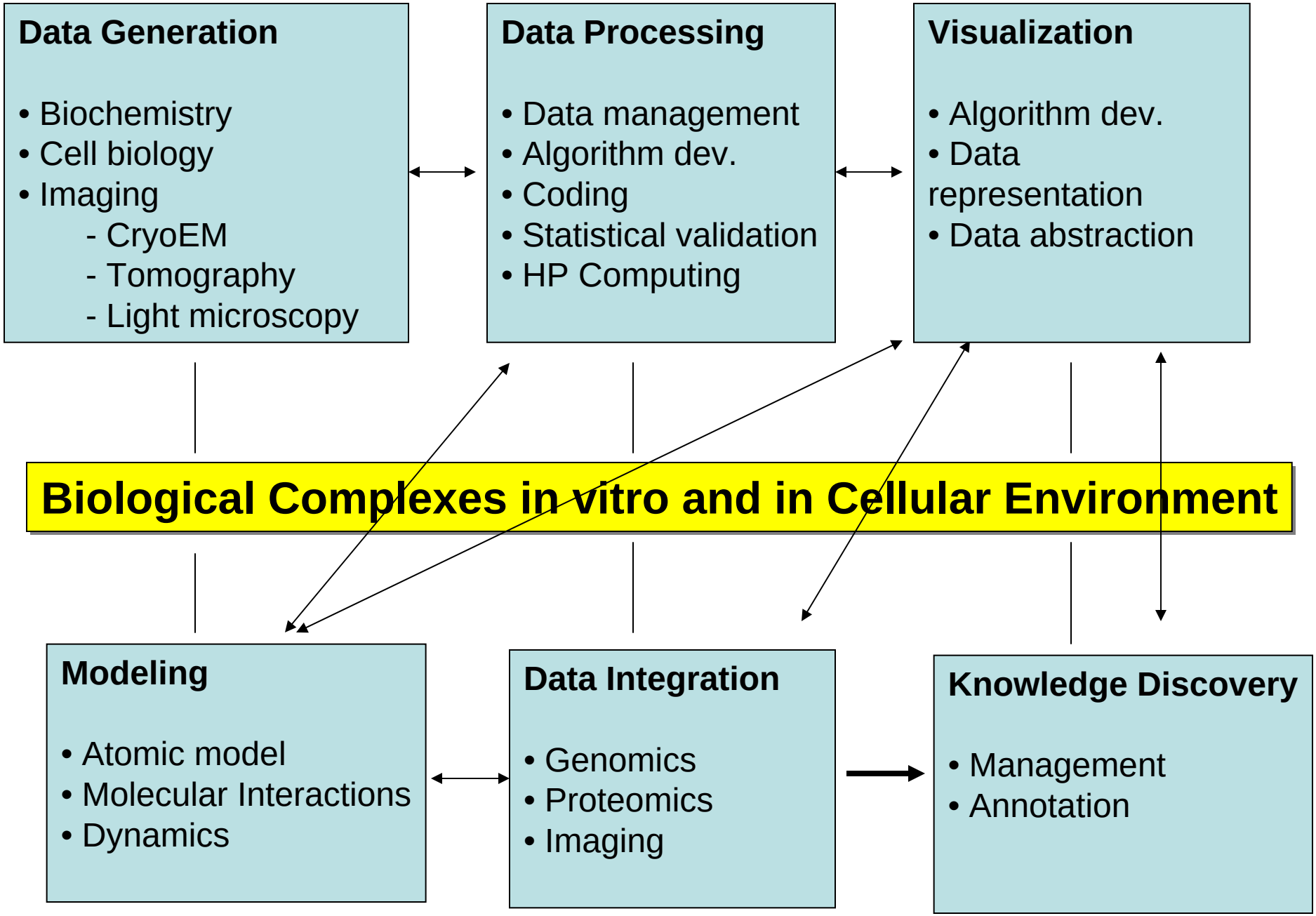
Engineering



Biological Science



Integrated Quantitative Biomedicine


Data Generation
• Biochemistry
• Cell biology
• Imaging
- CryoEM
- Tomography
- Light microscopy
Data Processing
• Data management
• Algorithm dev.
• Coding
• Statistical validation
• HP Computing
Visualization
• Algorithm dev.
• Data representation
• Data abstraction
Biological Complexes in vitro and in Cellular Environment
Modeling
• Atomic model
• Molecular Interactions
• Dynamics
Data Integration
• Genomics
• Proteomics
• Imaging
Knowledge Discovery
• Management
• Annotation

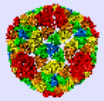
Visualizing Data

EMBL-EBI

Macromolecular Structure Database

home > services > em-atlas pages

Entry EMD-1060



Title: Rice dwarf virus

Authors: Zhou ZH, Baker ML, Jiang W, Dougherty M, Jakana J, Dong G, Lu G, Chiu W.

Aggregation State: icosahedral, (resolution 6.8 Angstroms)

- Summary
- Visualisation
- Sample

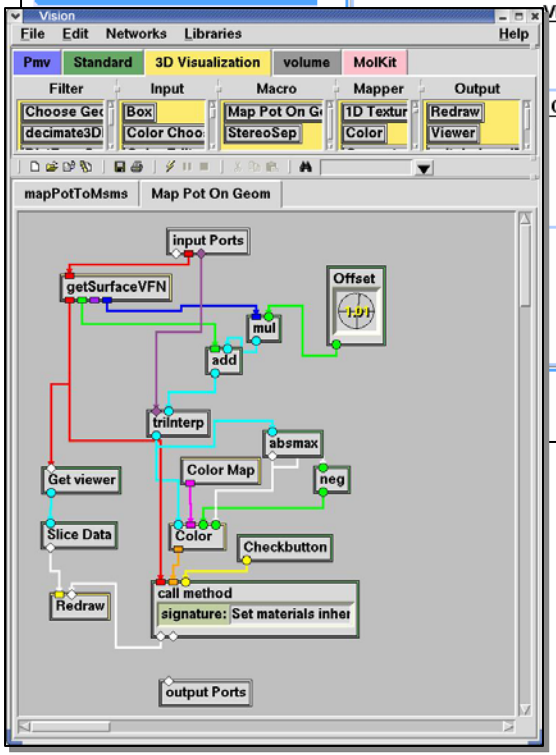
Experiment

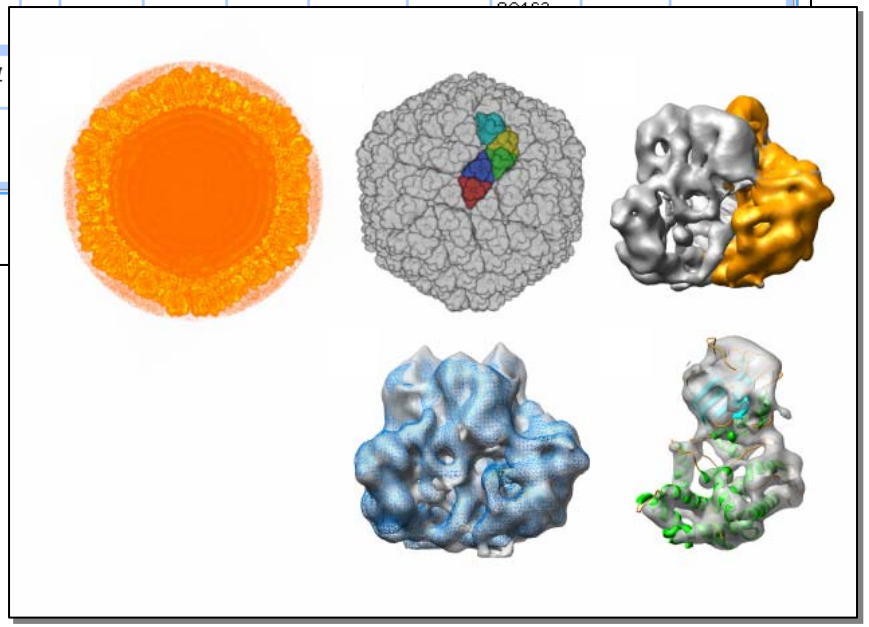
Sample Preparation:	ph:	Sample Conc.:	Details:	Staining:	Sample Support Details:

Vitrification:	Cryogen Name:	Humidity:	Temp.:	Instr.:	Method:	Time Resolved:	Details:
	ETHANE	27	111	manual			

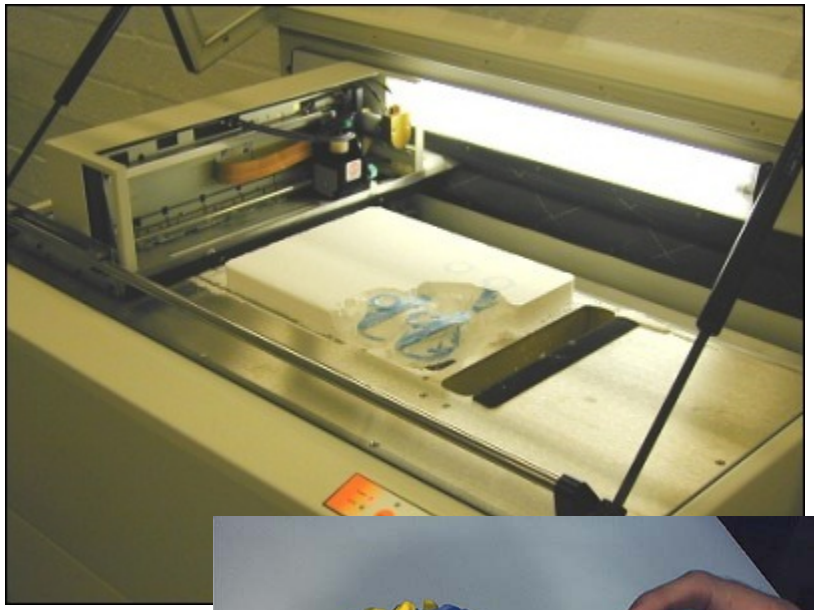
Conditions:	Microscope:	Voltage:	Illumination Mode:	Imaging Mode:	Cs:	Defocus Min:	Defocus Max:	Nominal Mag:	Calibrated Mag:	Electron Source:	Detector:	Detector distance:	Astigmatism:
	OTHER	400	FLOOD BEAM	BRIGHT FIELD	4.1	0.3	2.2	50000	49495	LAB6	Kodak	200	

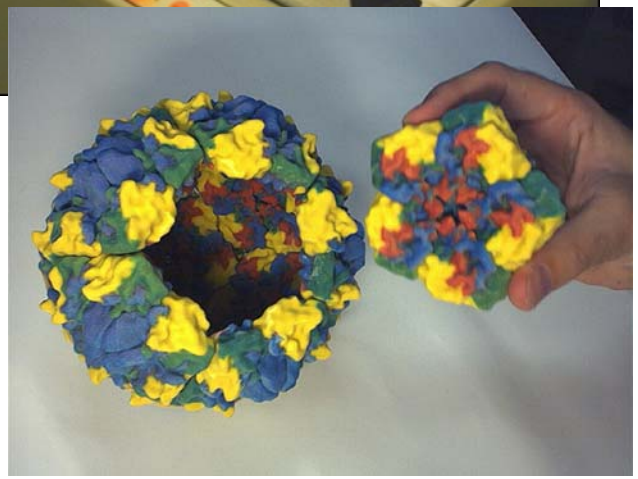
Specimen Holder:	Holder Model:	Tilt Min:	Tilt Max:	Energy Filter:
Gatan	GATAN LIQUID NITROGEN			



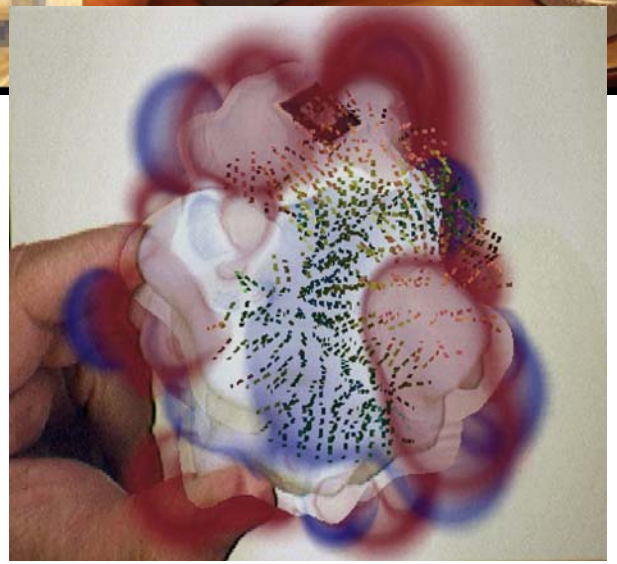


Tangible Models

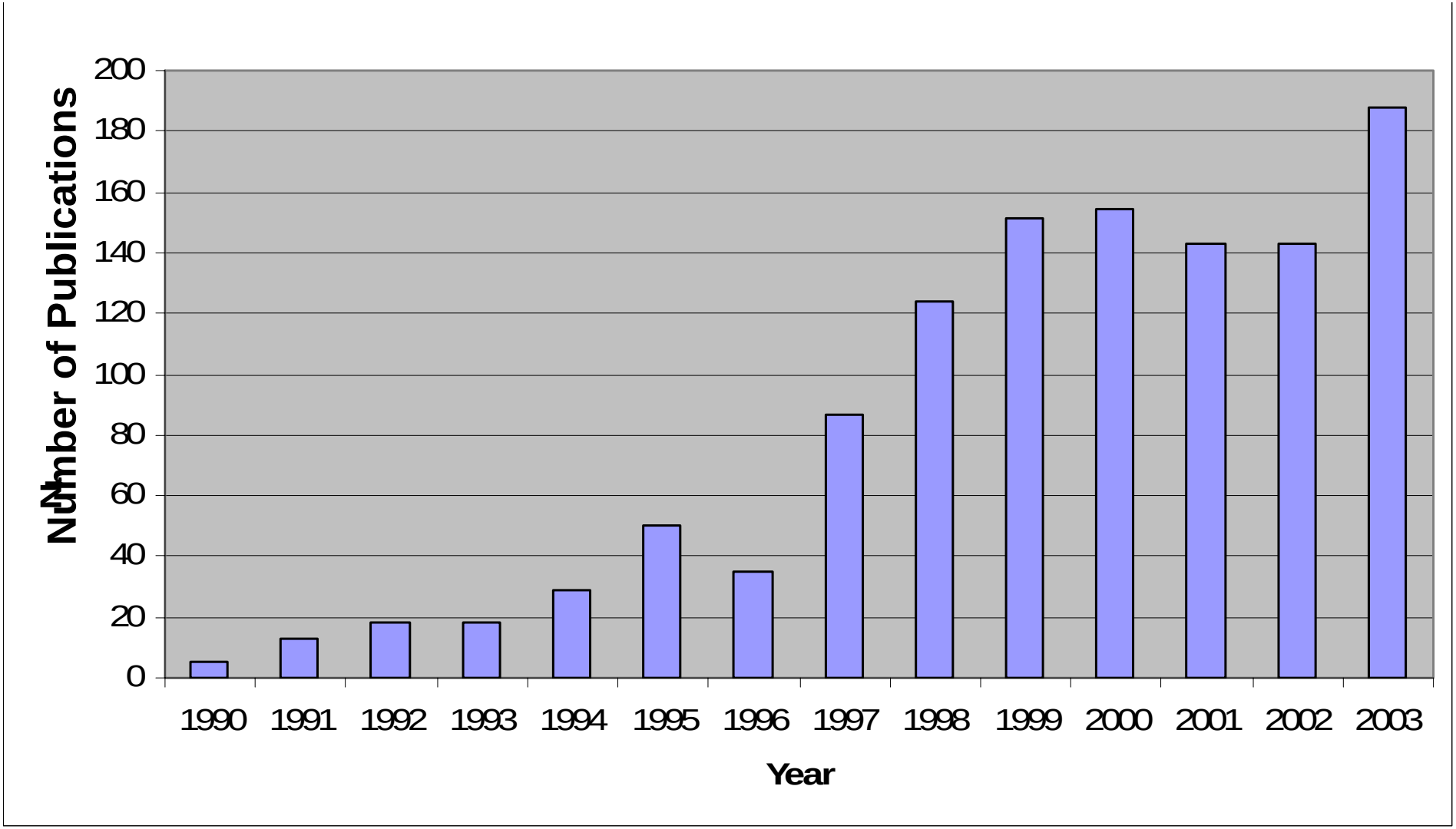




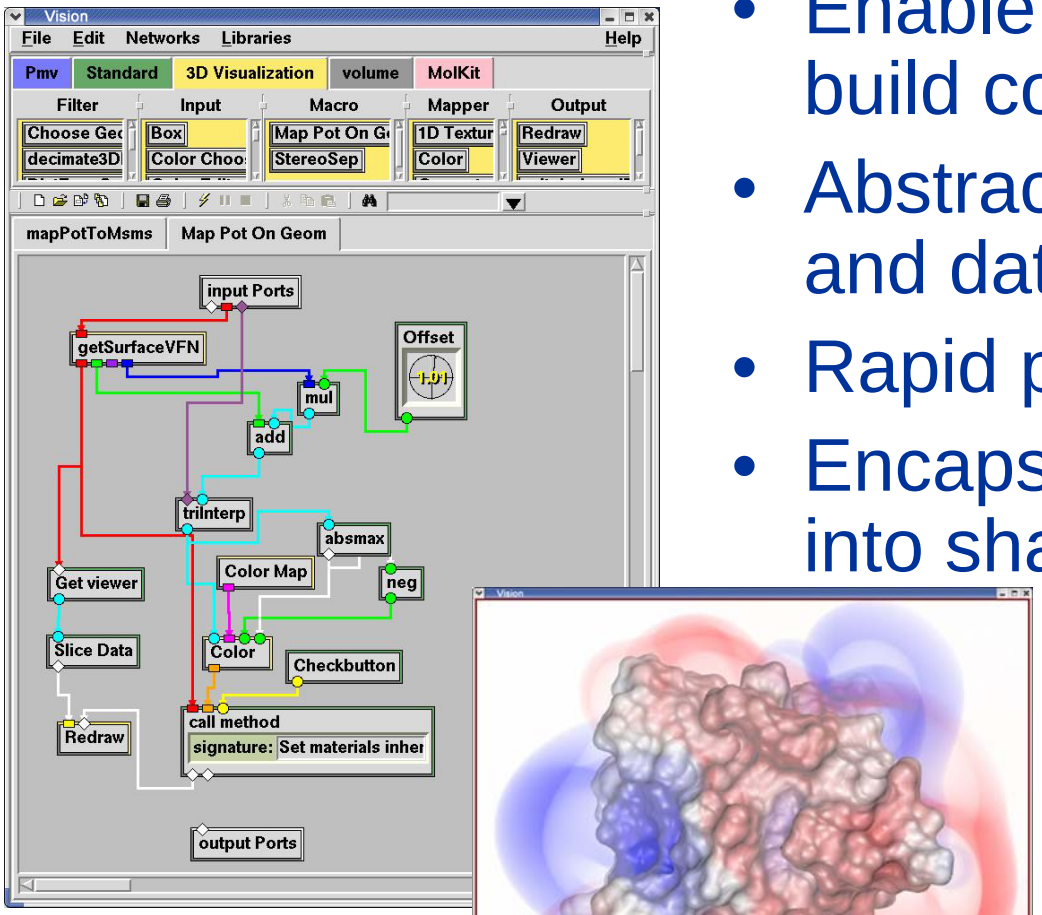
Augmented Reality



CryoEM and Cryomicroscopy Publications

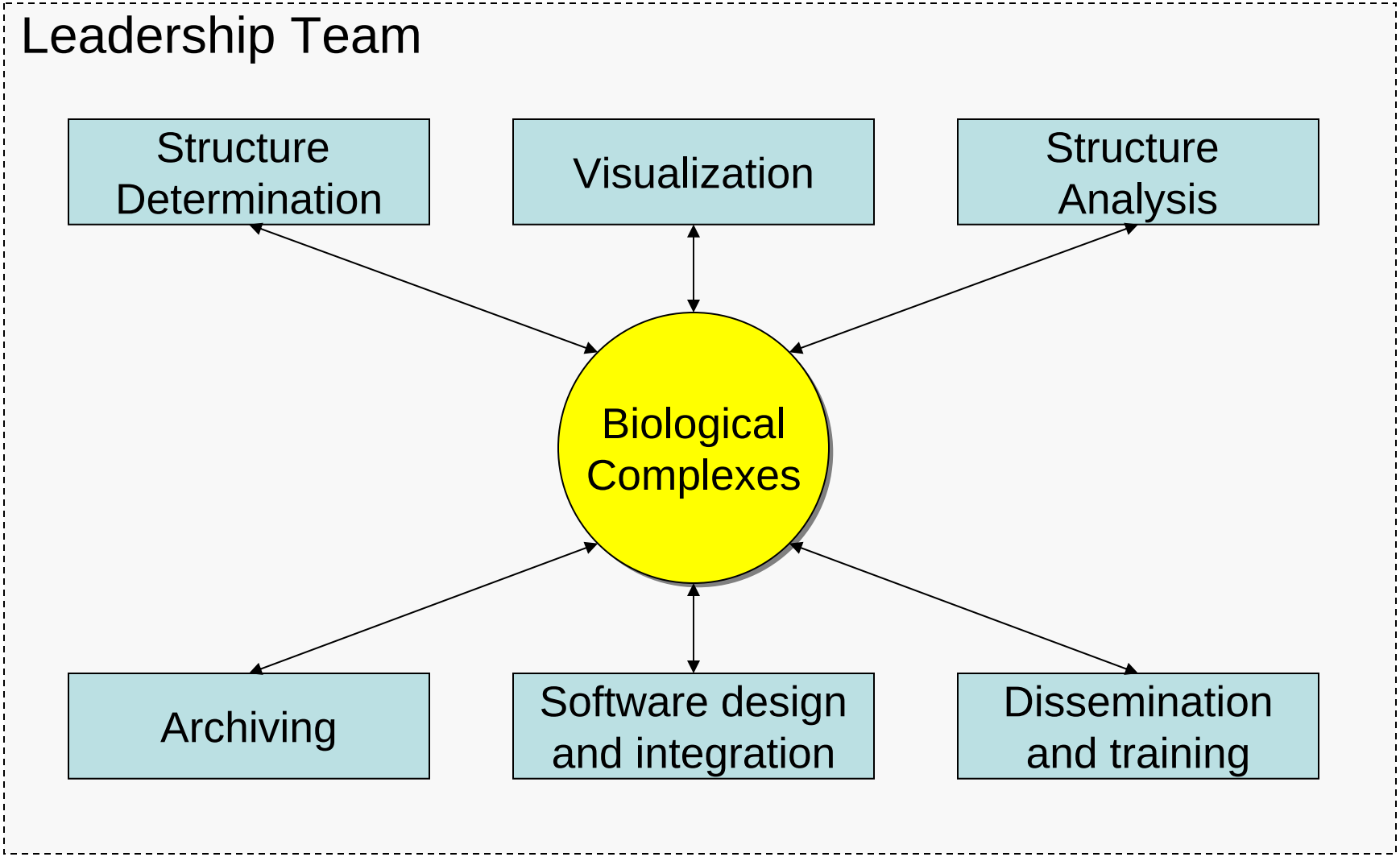


Vision: Visual Programming



- Enable non-programmers to build computational networks
- Abstract programming syntax and data structures
- Rapid prototyping
- Encapsulation of basic tasks into shareable computational

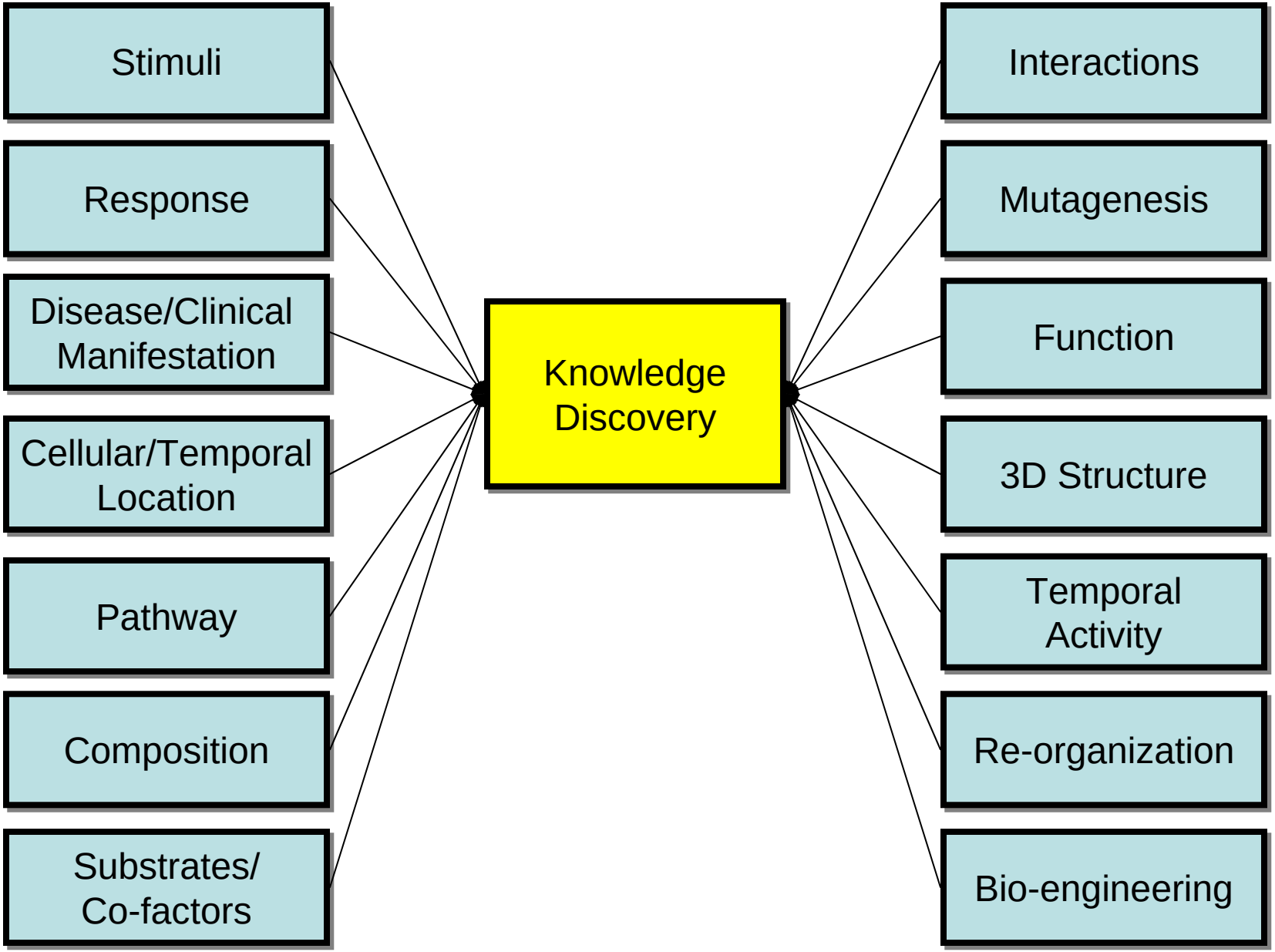
C²BC Vision: Organization



Discovery in Biological Processes

- How do we identify complexes (transient and stable) in a pathway?
- How do we isolate these complexes?
- Where in the cell are these complexes?
- How do we determine the structure of these large complexes?
- How do the components interact with each other?
- Can these complexes be altered/re-repurposed?
- How does this relate to disease?
- How do you visualize this data?
- Is there a common language to discuss the data?
- How do you integrate, disseminate and archive this information?

Discovery in Biological Processes



Discovery in Biological Processes

Stimuli	What causes/antagonizes the system?	Interactions
	How do the components interact?	
Response	What is the physiological response?	Mutagenesis
	How do mutations effect the process?	
Disease/Clinical Manifestation	How does this effect the system?	Function
	What are the functions of the components?	
Cellular/Temporal Location	Where and when do these events happen?	3D Structure
	What are the structures of the components?	
Composition	What is involved in the process?	Re-assembly
	Do the component structures change?	
Pathway	What is the order of events?	Temporal Activity
	What is the time-scale for activity?	
Substrates/ Co-factors	What is needed in the process?	Bio-engineering
	Can the process/components be re-purposed?	

- Platform and programs
- Project leaders are not tenured faculty
- Grow incrementally
- Core member and assoicated member
- Support postdocts and generate new centers
- Pilot projects
- Abstracts
- Proposals
- Papers